

AD-A282 928



(1)

FINAL 1 Dec 93 TO 28 Feb 94

VISUAL MOTION PERCEPTION AND VISUAL INFORMATION  
PROCESSING

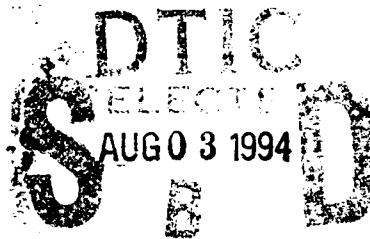
F49620-94-1-0073

61102F

2313/AS

Dr George Sperling

Univ of California, Irvine  
Dept of Cognitive Science  
School of Social Sciences  
Irvine, CA 92717



AFOSR-TR 94 0470

Dr John F. Tangney  
AFOSR/NL  
110 DUNCAN AVE SUITE B115  
BOLLING AFB DC 20332-0001

DISTRIBUTION STATEMENT A

Approved for public release  
Distribution Unlimited

218X 94-24544



This final progress report summarizes the results of a study that successfully determined the functional architecture of visual motion perception in the sense of defining the mechanisms involved and the relations between them. It was proved that visual motion is computed by two neural systems: primitive motion-energy extraction (e.g., Reichardt detector) and higher-level feature tracking. A psychophysical pedestal paradigm was used to exclude the feature-tracking process and thereby to obtain pure measures of motion-energy extraction. Motion energy extraction was found to be exclusively monocular, fast (cutoff frequency is 12 Hz) and sensitive (can utilize 0.2% contrast), "bottom-up", and to operate on both luminance (first-order) and contrast (second-order) motion stimuli. Motion feature tracking was found to operate interocularly as well as monocularly, have a cutoff frequency of 3 Hz, and to be both bottom up (it computes motion from luminance, contrast, depth, motion-motion, flicker and other type of stimuli) and top-down (e.g., attentional states influence what appears to move). The full report is appended.

(U)

(U)

(U)

(U)

94 8 02 193 DTIC QUALITY INSPECTED 1

AFOSR-TR- 94 0470

80044

Sperling: Visual Motion Perception and Visual Information Processing

FINAL REPORT AFOSR Award Number F49620-94-1-0073

December 1, 1993 to February 28, 1994 (includes unfunded extention to 31may94)

ABSTRACT

This final progress report summaries the results of a study that successfully determined the functional architecture of visual motion perception in the sense of defining the mechanisms involved and the relations between them. It was proved that visual motion is computed by two neural systems: primitive motion-energy extraction (e.g., Reichardt detector) and higher-level feature tracking. A psychophysical pedestal paradigm was used to exclude the feature-tracking process and thereby to obtain pure measures of motion-energy extraction. Motion energy extraction was found to be exclusively monocular, fast (cutoff frequency is 12 Hz) and sensitive (can utilize 0.2% contrast), "bottom-up", and to operate on both luminance (first-order) and contrast (second-order) motion stimuli. Motion feature tracking was found to operate interocularly as well as monocularly, have a cutoff frequency of 3 Hz, and to be both bottom up (it computes motion from luminance, contrast, depth, motion-motion, flicker and other types of stimuli) and top-down (e.g., attentional states influence *what* appears to move). The full report is appended.

Approved for public release;  
distribution unlimited.

Accession For	
NTIS GRA&I	<input checked="" type="checkbox"/>
DTIC Add	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Jurisdiction	
By _____	
Distribution _____	
Availability _____	
Dist	
A-1	

## Sperling: Visual Motion Perception and Visual Information Processing

## FINAL REPORT AFOSR Award Number F49620-94-1-0073

December 1, 1993 to February 28, 1994 (includes unfunded extention to 31may94)

## Project: The Functional Architecture of Human Visual Motion Perception

Historically, visual motion perception has been a central problem in perceptual theory. On the one hand, motion appears to involve an early stage of pattern recognition (the "same" pattern must be located first here and then there); on the other hand, motion appears to invoke a unique perceptual experience quite different from pattern or shape perception.

Almost from the beginning of the experimental study of motion perception, it has been evident that more than one kind of computation is involved, and there has been a plethora of dual-process and multi-process motion theories. Some recent examples are short- versus long-range motion (1), motion-energy and Reichardt detectors (2) versus zero crossings (3) or gradients (4), first-order versus second-order motion (5), and so on. While there clearly is a kernel of truth underlying each of these dichotomies, there have been two pervasive problems: so far, there have not been operations that give pure measures of each proposed mechanism, nor has there been a clear distinction between the algorithm by which motion is computed and the preprocessing of the visual image prior to the point of motion computation.

Here, we offer a combination of two basic paradigms (pedestal and interocular displays) plus several subsidiary paradigms (stimulus superpositions with varying phases and directions, stimulus mixtures, and attentional manipulations) that offer a clear indication of the motion algorithms and yield surprising insights into the image transformations involved. The pedestal paradigm will be self-evident, but its ultimate significance is that it offers a litmus test for a Reichardt (or the equivalent motion-energy) algorithm, so we briefly review these first.

## Reichardt (and Motion Energy) Models

Computational theories of motion perception date from Reichardt's model for insect vision (6), which was adapted for human perception by van Santen & Sperling (2). A Reichardt Detector consists of two mirror-image subunits (e.g., "Left" and "Right") tuned to opposite directions of motion (Fig. 1a). Subunit R multiplies the signal at spatial location A with the *delayed* signal at a rightward adjacent spatial location B. Subunit L multiplies signal at spatial location B with the *delayed* signal at spatial location A. The output of each subunit is integrated for a period of time and the direction of movement is indicated by the sign of the difference between the subunit outputs (6). Subsequent theories (7,8) of motion perception involving, essentially, Fourier analysis of the x,y,t motion stimulus to compute "motion energy" were shown to be computationally

equivalent to the Reichardt model (8).

-----  
Insert Figure 1 here.  
-----

Reichardt motion analysis is most naturally applied directly to drifting modulations of *luminance* that typically represent rigidly moving objects. Because it can be applied directly to a luminance signal, it is called 1st-order motion extraction (5). However, Chubb and Sperling (9) demonstrated clear motion perception in a broad classes of drift-balanced and microbalanced stimuli that were constructed of drifting modulations of contrast, spatial frequency, texture type, or flicker (see also 10). Such stimuli are said to activate 2nd-order motion mechanisms (5) because their motion is invisible to Reichardt or motion-energy detectors. Chubb and Sperling (9) noted that grossly nonlinear preprocessing (e.g., absolute value or square-law rectification) prior to a Reichardt detector could expose the latent motion in driftbalanced and microbalanced stimuli. Because first-order motion can be observed with smaller-size stimuli than is the case for second-order motion, it is generally believed that "short-range" and "first-order" are the the co-defining characteristics of one of the motion systems. We shall soon see the situation is more complicated than this.

Van Santen & Sperling (2) proved two extremely useful properties of Reichardt detectors (and of the equivalent motion-energy systems): (1) *Pseudo-linearity*: When a stimulus is composed of several component sine waves with different temporal frequencies, the detector's response to the sum is the sum of the responses to individual inputs (pseudo-linearity because linearity holds only for sines of *different* temporal frequencies). (2) *Static displays are ignored*: The output to any sinusoid of zero temporal frequency --a stationary pattern--is zero. From (1) and (2), it follows that adding a stationary sine (temporal frequency is zero, therefore output is zero) to any moving pattern (*moving* means temporal frequency is nonzero) does not change the output of a Reichardt detector to the moving stimulus.

#### The Pedestal Test

We exploit the pseudo-linearity of Reichardt detectors by creating compound stimuli consisting of a stationary sine (the pedestal, Figs. 1b & 1e) plus a linearly-moving sine grating (the test, Figs. 1c & 1f). The peaks and valleys of the compound stimulus oscillate back and forth (Figs. 1d & 1g). Nevertheless, the output of a Reichardt detector is exactly the same for the pedestal+test stimulus as for the test alone. In practice, nonlinearities of human vision before and after the movement computation require that the amplitudes of these sine stimuli be small (e.g., less than about 5 percent modulation depth) in order for the pedestal predictions to hold exactly. The question is: how do human observers perceive the compound stimulus? Do they track the peaks (which implies a feature tracking mechanism) or do they perceive the linear motion of the test stimulus?

The pedestal-plus-test stimulus is defined in the x,t domain. Consider the equivalent pedestal-plus-test defined in the x,y domain, in which the pedestal is a vertical grating and the test is a slanted grating. Such an x,y texture is displayed (at high contrast) in Fig. 1g, and the answer is obvious. Observers do not directly perceive the component gratings; they perceive primarily back-and-forth oscillation, and an (apparent) amplitude modulation. In these illustrations, the

stationary spatial sine wave pedestal has twice the amplitude of the linearly moving stimulus; it produces a back-and-forth phase oscillation equal to 1/6 of the spatial cycle (Fig. 1d), and this phase oscillation is what is perceived (11).

By a pedestalled stimulus, we refer to a pedestal-plus-motion stimulus with a 2:1 pedestal: test amplitude ratio. We conducted formal experiments to determine how pedestalled tests are perceived in the x,t motion domain. To reiterate: the Reichardt model predicts that motion direction extraction should be completely unaffected by the pedestal. Therefore, we first determine each subject's threshold amplitude for the discrimination of a leftward from a rightward moving pure sinewave grating. We then add a pedestal with twice this measured threshold amplitude. If the judgment were based on the output of a Reichardt detector, we expect the subject's accuracy of left versus right judgments to be the same with and without the pedestal. On the other hand, if the motion direction computation were based on stimulus features (peaks, valleys, zero-crossings, etc), the pedestalled motion would appear oscillatory, and it would be impossible for subjects to judge motion direction of the test.

#### Stimuli

We constructed four different types of motion stimuli: a luminance grating (Fig. 1e), a contrast grating (Fig. 1h), a depth grating without any monocular motion cue (Fig. 1k), and a motion-defined moving grating (Fig. 1l). Except for the luminance stimulus, only the pattern of modulation, not the stimulus itself moves (either to the left or to the right).

*Luminance grating (12).* The luminance stimulus is the sort of first-order motion stimulus, a rigidly translating sinewave pattern (Fig. 1b and 1f) from which traditional motion psychophysics has evolved.

*Contrast grating (13).* The contrast grating is a pure second-order stimulus: Its expected luminance is the same everywhere; its motion cannot be determined by Reichardt detectors. However, an initial stage of spatial filtering, followed by a nonlinearity such as fullwave rectification (e.g., absolute value or squaring) can expose the contrast grating's motion to standard motion analysis (e.g., Reichardt detectors; see Chubb & Sperling, 1989b).

*Depth grating (14).* The dynamic stereo-depth grating is created from stereo views of left- and right-half images. It appears in depth as a surface whose distance from the observer varies, as illustrated. The grating (and its depth) exist only as a space-varying correlation between the pixels in the left- and right-eye images; each monocular image is completely homogeneous without any hint of a grating, and successive images are uncorrelated.

*Motion-defined motion (15).* The motion-defined grating consists of dots that make step jumps in successive frames. The proportion of upward versus downward jumping dots varies sinusoidally from left-to-right. To perceive the movement of the motion-defined grating requires (1) computing the direction of motion of the dots, and (2) noting that the sinewave pattern of dot-motion moves with time. This kind of "motion-from-motion" (16) seems to suggest a hierarchical organization of motion detectors.

### Procedure

Within an experimental session (17), only one type of stimulus was presented, but various temporal frequencies were mixed. Subjects initiated a trial by pressing a key. A fixation point immediately appeared in the center of the display and remained on throughout the trial; 0.5 sec later, the stimulus was presented. The stimulus was always presented for a full temporal cycle starting with a random phase. To remove locational cues, the first and last frames were always identical. The subjects' task was to indicate by a key press which one of two possible motion directions was perceived and to give a confidence rating ranging from 0 (totally uncertain) to 5 (absolutely sure).

Initially, subjects' motion-direction discrimination thresholds were measured without pedestals. The method of constant stimuli was used to generate psychometric functions for a set of temporal frequencies for each motion stimulus type. At least 100 observations were made for each subject at the five points (determined by preliminary observations) that best defined the psychometric function (probability correct versus modulation amplitude). For a given motion stimulus of type  $s$  and temporal frequency  $f$ , we defined the subject's threshold as the amplitude  $m_{75}(s, f)$ , corresponding to the 75% correct point on the psychometric function. In the subsequent pedestal test, the amplitude of the motion stimulus was always set at  $m_{75}(s, f)$ . On each trial, the modulation amplitude of the pedestal was randomly set either at 0 or at  $2m_{75}(s, f)$ . Within a session, all temporal frequencies and pedestal amplitudes were mixed. We compared every subject's performance with and without the pedestal in the same session. Two male subjects with corrected-to-normal vision served as subjects for the data reported here.

### Results

Both subjects clearly perceived apparent motion in all the motion-stimulus-alone conditions when the sine amplitude was sufficient, and on the whole, produced quite similar data. The quantitative results of one are summarized in Fig. 1m. (A) The temporal tuning functions for all the motion types show typical lowpass filter characteristics (curves slope down to the right) within the temporal frequency range we tested (0.94 to 15.0 Hz). (B) The temporal tuning functions can be divided into two groups: luminance grating and contrast grating as one group (upper curves, Fig. 1m), depth-defined grating and motion-defined grating as another group (lower set of curves). Within each group, the shape of the temporal tuning functions is remarkably similar. (C) The presence of a 2x pedestal had absolutely no effect on subjects' performances in the luminance and contrast modulation conditions (18), but it reduced performance to merely chance-guessing levels with the depth-defined and motion-defined gratings. For pedestalled depth and motion-motion stimuli, subjects reported that they perceived only back-and-forth motion, and could not judge the direction of (apparently invisible) coherent motion (19).

These results clearly indicate that there are two qualitatively different motion extraction mechanisms. One mechanism utilizes only the motion energy computation and has a much higher cutoff (12 Hz) in its temporal sensitivity characteristics. It subserves both luminance (first-order) and contrast (second-order) stimuli. Interestingly, the contrast-motion system has the same temporal frequency characteristics as the luminance-motion system, despite frequent speculation that the second order system is "slower" than the first order system (20). The second mechanism is slower, but can detect motion in stimuli that are invisible to the first mechanism.

#### Four Confirming Procedures

To further understand these results, we report briefly four subsequent procedures (21).

(A) We superimpose (linearly add) pedestalled luminance and contrast stimuli moving in the same direction or in opposite directions (each with its own pedestal). Adding equal strength motion stimuli in opposite directions cancels any perceived motion. The perceived motion strength of stimuli moving in the same direction is given by probability summation of the strengths of the two component stimuli; there is no dependence on the relative phases of the two stimuli. If the two kinds of stimuli were combined prior to the motion computation, the *sign* of the combination would depend on the relative phase of the components; indeed, stimuli of the same frequency moving in the same direction but with opposite phases could cancel all perceived motion. The absence of any phase dependence means that luminance- and contrast-motion strengths are first computed separately; then, the two motion strengths are combined.

(B) A pedestalled stimulus was created with only four frames per cycle, successive frames being separated by 90 deg. Motion in this stimulus is perceived as well as in a continuously sampled stimulus (Fig. 1f versus Fig. 2). However, directing successive frames alternately into left and right eyes absolutely destroyed our subjects' ability to perceive the direction of motion. With pedestalled stimuli, the direction of motion can be computed only monocularly, not interocularly.

---

Insert Figure 2 here.

---

(C) On the other hand, presenting successive frames of the motion-from-motion stimulus (with *no* pedestal) to alternate eyes (interocular presentation, Fig. 2) only slightly increases threshold for motion-direction discrimination relative to a monocular presentation. This indicates that the motion-from-motion computation is inherently binocular.

(D) Consider the display of a luminance sinusoid with successive frames separated by 90 deg (Fig. 2). It can be viewed either monocularly (all frame in same eye) or interocularly (successive frames in alternate eyes). Converting from monocular to interocular presentation raises the contrast threshold (at low frequencies) by a factor of 12 (to 2%) and changes the frequency cutoff from 12 to 3 Hz, exactly like that of the depth and motion-from-motion stimuli (Fig. 1m). This shows that the interocular luminance grating is perceived by the feature-tracking mechanism; this mechanism exhibits exactly the same frequency cutoff when it detects motion luminance stimuli as it does when it detects motion in depth and motion-from-motion stimuli.

A consequence of the above is that the motion of an apparently simple stimulus, such as a drifting luminance grating, is computed by all three systems: The monocular luminance system, which is fast and sensitive; the binocular feature tracking system which is slow and less sensitive; and at a double frequency, the monocular fullwave-contrast system (which is relatively insensitive to this kind of stimulus). The drifting grating, which is regarded as a universal tool for visual psychophysics, turns out to be not a particularly useful tool for discriminating among motion mechanisms.

### Selective attention affects the direction of perceived motion

Suppose that the binocular mechanism were a feature-tracking mechanism. Then it would track whatever features are dominant in successive stimuli, even when the successive stimuli are composed of entirely different materials. We tested this prediction by alternating depth gratings with texture gratings. Frames 1,3, and 5 consisted of a depth grating; frames 2 and 4 consisted of a texture grating (Fig. 3). In depth gratings, all subjects naturally attend to the near-appearing peaks. In our texture grating, in successive sessions, subjects were asked to attend either to the right-slanting higher-spatial frequency grating or to the left-slanting lower spatial frequency grating. The display was arranged so that if one texture feature were dominant, the display would appear to move to the right; if the other were dominant, it would appear to move to the left. From trial to trial, these relations were reversed randomly. If an observer were to selectively track only the grating feature, or only the depth feature, the display would be completely ambiguous. To perceive unambiguous motion, the observer must bind two features, i.e., perceive movement from the attended depth feature (near) to the attended texture feature (e.g., left-slanted coarse texture).

-----  
Insert Figure 3 here.  
-----

In formal experiments, sequences of five successive stimuli (A,B,C,D,A; Fig. 3) were presented at a frequency of 2.5 Hz so that an entire display was completed in 500 msec. Two subjects consistently perceived motion in the direction corresponding to the attended feature in more than 95% of trials (chance equals 50%). Thus, the same stimulus (Fig. 3) was perceived as moving to the right when observers attended to fine texture and as moving to the left when they attended to the coarse texture. This indicates that not only stimulus properties but also attention determines what features are tracked (22).

The influence on perceived motion by selective attention is inherently a top-down process. Verbal instructions to the subject prior to the trial are processed at a cognitive level. The output of this high-level process is used to control a low-level filter, which controls the input to the feature-tracking mechanism by selectively admitting either coarse or fine textures. The low cutoff frequency of the feature-tracking system, about 3 Hz, suggests that the shortest period within which attention can be moved is about 0.33 sec, a period that roughly corresponds with the dynamics of shifts of visual attention that are measured in quite different paradigms (23). These results indicate that the feature-tracking mechanism is affected by both top-down attentional processes and by automatic bottom-up processes, each of which contributes to the strength of the tracked features.

### Discussion and Conclusions

*The fast, monocular system.* The results reported above are embodied in the functional flowchart of Fig. 4. The left side of Fig. 4 shows the fast monocular system: separate Reichardt detectors for inputs from each eye, and separate detectors for luminance (first-order) and contrast (second-order) stimuli within the eye. These detectors are fast (cutoff freq = 12 Hz), sensitive (can detect contrasts of 0.2%), and the outputs are combined after motion is computed. Survival of pedestal tests indicates the fast mechanisms use Reichardt (or equivalent) mechanisms for all

stimuli being processed; failure to survive interocular presentations indicates purely monocular mechanisms.

-----  
Insert Figure 4 here.  
-----

*Feature tracking.* The right-hand side of Fig. 4 represents the feature-tracking system. It is almost as efficient in detecting motion in interocular as in monocular displays; i.e., it is inherently binocular. Like the fast mechanism, it can compute motion of luminance and contrast stimuli (although in a more restricted speed range and with much less sensitivity). It can also compute motion for stimuli that are invisible to the fast mechanism if they have readily identifiable features (such as moving depth gratings and motion-from-motion displays). To solve the motion-from-motion displays requires input of the direction-of-motion feature from the monocular system. For all stimuli, feature tracking exhibits a characteristic 3 Hz cutoff. Attention can determine which of two competing features is tracked, this top-down control of the inputs to the feature tracking computation is indicated by the arrow from Cognitive Processing to a feature-weighting component.

The mechanism of motion detection in feature tracking system has not determined. This is because feature specification is inherently coarsely quantized (a feature is either present or absent). A Reichardt computation fails with a coarsely quantized pedestalled stimulus. Therefore, we cannot say whether the failure of pedestalled motion to survive interocular manipulations is caused by a Reichardt mechanism confronted with a too-coarsely quantized (or too-noisy) stimulus or whether feature tracking uses an entirely different algorithm. From a biological point of view, it seems plausible that all motion computations would use a similar algorithm, perhaps embodied in a common patch of genetic code, and that only prior transformations and spatio-temporal parameters would distinguish the two levels of computation.

*Higher processes, methodology.* The current model deals with motion-direction discrimination. The outputs, especially of the fast first-order Reichardt detectors, have been proposed as the inputs to perceptual processes that compute velocity (24), 3D structure from motion (25), and other useful properties. The pedestal test provides a useful means for distinguishing between classes of motion extraction mechanisms, and we anticipate its application in these and other stimulus domains.

#### REFERENCES AND NOTES

1. O. Braddick, *Vision Research*, **14**, 519 (1974); A. Pantle and L. Picciano L. *Science*, **193**, 500 (1976); G. Mather, P. Cavanagh and S.M. Anstis, *Perception*, **14**, 163 (1985).
2. J. P. H. van Santen and G. Sperling, *J. Opt. Soc. Am.* **1**, 451 (1984).
3. D. Marr and S. Ullman, *Proc. R. Soc. London Sect B* **211**, 151 (1981);
4. E.H. Adelson, and J.R. Bergen, in *Motion: Representation and Analysis* (IEEE Workshop Proceedings), 151 (1986).
5. P. Cavanagh and G. Mather, *Spatial Vision* **4**, 103 (1989); C. Chubb and G. Sperling, *Pro. of*

*the Nat. Acad. of Sci. U.S.A.* **86**, 2985 (1989); H.R. Wilson, V.P. Ferrera and C. Yo, *Visual Neuroscience*, **9**, 79 (1992); J.C. Boulton and C.L. Baker, *Vision Research*, **33**, 2013 (1993).

6. W. Reichardt, *W. Zeitschrift Naturforsch.*, **12b**, 447 (1957); W. Reichardt, in *Sensory Communication*, W. A. Rosenblith, Ed. (Wiley, New York, 1961);
7. E. H. Adelson and J. R. Bergen, *J. Opt. Soc. Am.* **2**, 284 (1985); A. B. Watson and A. J. Ahumada, *J. Opt. Soc. Am.* **2**, 322 (1985);
8. J. P. H. van Santen and G. Sperling, *J. Opt. Soc. Am.* **2**, 300 (1985).
9. C. Chubb and G. Sperling, *Proceedings: Workshop on Visual Motion*, Irvine, CA. IEEE Computer Society Press, 126 (1989b); C. Chubb and G. Sperling, *J. of Math. Psych.* **35**, 411 (1991).
10. V. S. Ramachandran, M. V. Rau and T. R. Vidyasagar, *Vision Research* **13**, 1399 (1973); A. M. M. Lelkens and J. J. Koenderink, *ibid* **24**, 1083 (1984); K. Turano and A. Pantle, *ibid* **29**, 207 (1989); P. Cavanagh, M. Arguin and M. von Grlnau, *ibid* **29**, 1197 (1989); J. D. Victor and M. M. Conte, *ibid* **30**, 289 (1989); G. Mather, *ibid* **31**, 161 (1991). A.M. Derrington & D.R. Badcock, D.R., *ibid*, **25**, 1869 (1985).
11. Zero crossings, the motion cue proposed by Marr and Ullman (see reference 3), follow the same back-and-forth path as do peaks or valleys.
12. Let  $x$  denote the horizontal spatial coordinate,  $y$  denote the vertical spatial coordinate. The luminance grating is defined as

$$L_1(x, y, t) = L_0[1.0 + m(1, f_1) \sin(2\pi(\alpha_1 x + f_1 t))]$$

where  $L_0 = 115 \text{ cd/m}^2$ ,  $\alpha_1 = 2.55 \text{ cpd}$ , and  $f_1 = 0.94, 1.88, 3.75, 7.50, 15.0 \text{ Hz}$ . The grating extends 3.13 degrees horizontally and 1.57 degrees vertically.

13. The contrast modulation grating is defined as:

$$L_2(x, y, t) = L_0[1 + R(x, y)(0.5 + m(2, f_2) \sin(2\pi(\alpha_2 x + f_2 t)))]$$

where  $L_0 = 115 \text{ cd/m}^2$ ,  $\alpha_2 = 1.28 \text{ cpd}$ ,  $f_2 = 0.94, 1.88, 3.75, 7.50, 15.0 \text{ Hz}$ , and  $R(x, y)$  is a random number which assumes value +1 and -1 with equal probability. The grating has the same size as the luminance grating.

14. The depth grating was made of white ( $129.4 \text{ cd/m}^2$ ) random dots ( $0.733' \times 1.466'$ ) on gray background ( $105.1 \text{ cd/m}^2$ ) with horizontal disparity between left and right eyes defined as:

$$D(x, y, t) = 0.733' \text{Int}(m(3, f_3) \sin(2\pi(\alpha_3 y + f_3 t)))$$

where  $\alpha_3 = 1.28 \text{ cpd}$ ,  $f_3 = 0.94, 1.88, 3.75 \text{ Hz}$ .  $\text{Int}()$  is a function that takes real numbers as input and rounds them off to produce integer outputs. It is due to the fact that the pixels are discrete. The stimulus for each eye extends 2.94 degrees vertically and 1.47 degrees horizontally. In each frame, there is a 40% probability for a dot to be white and no correlation exists between frames.

15. The motion defined moving stimulus was also made of white dots ( $215.6 \text{ cd/m}^2$ ,  $5.84' \times 5.84'$ ) on gray background ( $115 \text{ cd/m}^2$ ). The probabilities for the random dots in a given column to move up  $P_u$  or down  $P_d$  are defined as:

$$P_u(x, y, t) = 0.5 + m(4, f_4) \sin(2\pi(\alpha_4 x + f_4 t))$$

$$P_d(x,y,t) = 1.0 - P_u$$

where  $\alpha_4=0.32cpd$ ,  $f_4 = 0.94, 1.88, 3.75$  Hz. There is a 40% probability for a dot to be white and a dot moves  $5.84'$  from one frame to the next one. The whole stimulus extends 7.04 degrees horizontally and 4.69 degrees vertically.

16. *Theta motion* (J.M. Zanker, *Vision Research*, **33**, 553 (1993).) Unfortunately, this report confounds motion-direction and "quantity of motion."
17. The experiment was controlled by an IBM 486PC compatible computer, driving a TrueVision AT-Vista video graphics adapter. The displays for the experiments were presented on a 60 Hz vertical retrace IKGAMI DM516A (20 inch diagonal) monochrome graphics monitor with a P4-type phosphor. The hardware setup can produce 12 bits (4096) distinct gray levels. The luminance of the monitor was  $.69cd/m^2$  when every pixel was given the lowest gray level and  $229.3cd/m^2$  when every pixel was given the greatest gray level. A lookup table had been generated with a psychophysical procedure which linearly divides the whole luminance range to 256 gray levels. We chose the background luminance to be that value which, when it is assumed by every pixel, produces  $0.5 * (229.3 + 0.69) = 115.0cd/m^2$ .
18. In a control experiment, we asked subjects to rate the visibility of the pedestals. The spatial frequency, temporal frequency, display duration and phase randomization are exactly the same as those used in the luminance and contrast modulation pedestal experiments. Here we removed the motion stimulus but left the pedestal unchanged. In all cases, the pedestals were visible. In fact, in a visibility scale from 0 to 5, the average visibility for pedestals is all above 2.0 and the pedestals are more visible at higher temporal frequencies. Thus, the failure of our pedestals to mask is independent of their (above threshold) perceived visibility.
19. One might argue that the subjects' inability to perform with pedestals is due to an early compressive nonlinearity in the motion extraction system. Our control experiments showed that subjects' performance increased when we increased the amplitude of the motion stimulus in the motion stimulus alone condition, even exceeding the 3x threshold maximum amplitude in the pedestal experiment. This suggests that we have not reached the saturation point of the system and a compressive nonlinearity can not account for the data. Nor does raising the amplitude of the entire 2:1 pedestal-plus-test stimulus admit improved performance.
20. A. M. Derrington, D. R. Badcock and G. B. Henning, *Vision Research* **33**, 1785 (1993).
21. Z.-L. Lu and G. Sperling, manuscript in preparation.
22. P. Cavanagh, *Science* **257**, 1563 (1992) reports that moving attention to track a feature causes a motion aftereffect opposed to the direction of supposed attentional movement. This is different than enhancing a stimulus feature by attention, although the same feature tracking mechanism might be involved.
23. A. Reeves and G. Sperling, *Psychological Review*, **93**, 180 (1986); E. Weichselgartner and G. Sperling, *Science*, **238**, 778 (1987).
24. D. Heeger, *J. Opt. Soc. Am.*, **4**, 1455 (1987);
25. G. Sperling, M. S. Landy, B. A. Dosher and M. E. Perkins, *J. Exptl. Psychol.: Human Percep. and Perf.*, **15**, 426 (1989); B. A. Dosher, M. S. Landy and G. Sperling, *Vision Research*, **29**, 1789 (1989).
26. This research was supported by Air Force Office of Scientific Research, Life Sciences, Visual Information Processing Program, Grant 91-0178. We thank Erik A. Blaser for his participation in the experiments.

### Figure Captions

Figure 1. (a) Reichardt motion detector (simplified). A and B indicate adjacent locations of visual receptive fields,  $\tau$  is a temporal delay,  $\times$  indicates multiplication, and  $-$  indicates subtraction. Outputs greater than zero indicate stimulus motion from A to B; outputs less than zero indicate stimulus motion from B to A. (b) A stationary sinewave (the pedestal). (c) A moving sinewave (the test stimulus). (d) Pedestalled motion: the sum of (b) and (c). The pedestal has twice the amplitude of the motion stimulus, which moves 1/8 of a spatial cycle to the right from one frame to the next. The dotted line indicates that the peaks in each frame follow a zigzag path and no coherent "net" motion direction exists for any mechanism that tracks the peak locations. (e,f,g) Pedestalled luminance-motion (first-order motion) (e) A freeze-frame snapshot of a luminance sinewave. Alternatively, interpreting the vertical axis as time (instead of vertical space) converts (e) into a space-time representation of a stationary luminance sinewave, an instantiation of (b). (f) A space-time representation of a moving luminance sinewave, an instantiation of (c). The vertical axis is time, the horizontal axis is space. (g) The sum of (e) and (f), a luminance grating moving over a pedestal, an instantiation of (d). (h,i,j) Pedestalled contrast-modulated motion (second-order motion). Similar to (e,f,g) except instead of a sinusoidal modulation of luminance, the contrast of a black-white noise field is sinusoidally modulated. (k) Representation of the appearance of the depth grating; the actual depth grating was composed of random black-white noise, the depth resulted from stereoscopic images. (l) Representation of the motion-from-motion stimulus. The arrows indicate the directions of motion of random dots; the pattern of motion modulation (up versus down) moves either to the left or to the right. (m) Experimentally measured threshold modulations for correct left-right motion discrimination versus temporal frequency (Hz) of a moving sinusoid. The axes are log scales.  $\circ$  indicates luminance (Fourier, FO) motion for either pedestalled or nonpedestalled stimuli (thresholds are identical);  $\Delta$  indicates contrast modulated (fullwave, FW) motion for either pedestalled or nonpedestalled stimuli;  $+$  indicates simple (nonpedestalled) sinusoidal depth (DP) stimuli;  $\times$  indicates simple sinusoidal motion-from-motion (MS) stimuli;  $?$  indicates simple sinusoidal interocular (IT) luminance stimuli. The curves have been vertically translated to expose their similarity in shape.

Figure 2. Schematic representation of interocular stimulus presentations. Stimuli are displayed only with spatial phase shifts of 90 deg relative to the initial display, indicated by the tick on top of the leftmost peak. At each eye, the stimulus sequence, indicated on the bottom, is ambiguous as to direction of motion.

Figure 3. An display for demonstrating attentional effects in the perceived direction of motion. A sequence of four consecutive displays is shown; each is displaced by 90 deg from the previous one. In the depth stimuli, the tracked features are always the near peaks (upper peaks in the panels). When the subject attends (tracks) the fine stripes, the perceived direction of motion between displays a-d is from left to right; attending to the coarse stripes yields right-to-left perceived motion. Duration of the entire display is only 0.50 sec, too fast to permit eye movements.

Figure 4. Functional architecture of the visual motion system. The left-half represents the fast monocular system; the right half represents the feature-tracking system. L, R indicate Left and Right eye signals, respectively; RD indicates Reichardt detector; TG indicates texture grabber (a spatial filter followed by fullwave rectification);  $\Sigma$  indicates (possibly complex) summation;  $\times$  represents multiplication, the differential weighting of features determined by attention (the arrow from "Cognitive Processes"); the central horizontal arrow represents the motion-feature input needed to solve motion-from-motion stimuli.

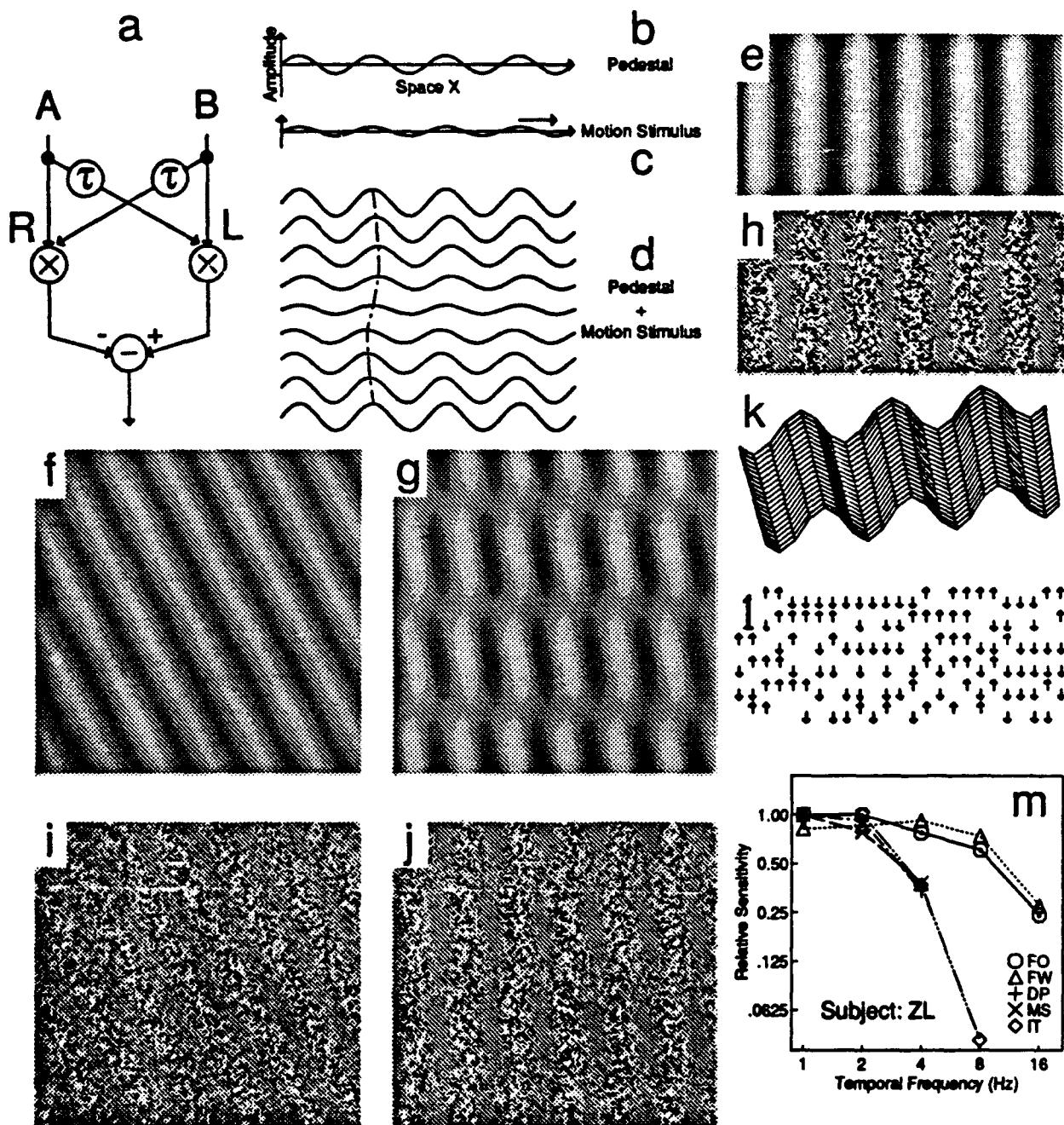


FIG. 1

Interocular Presentation

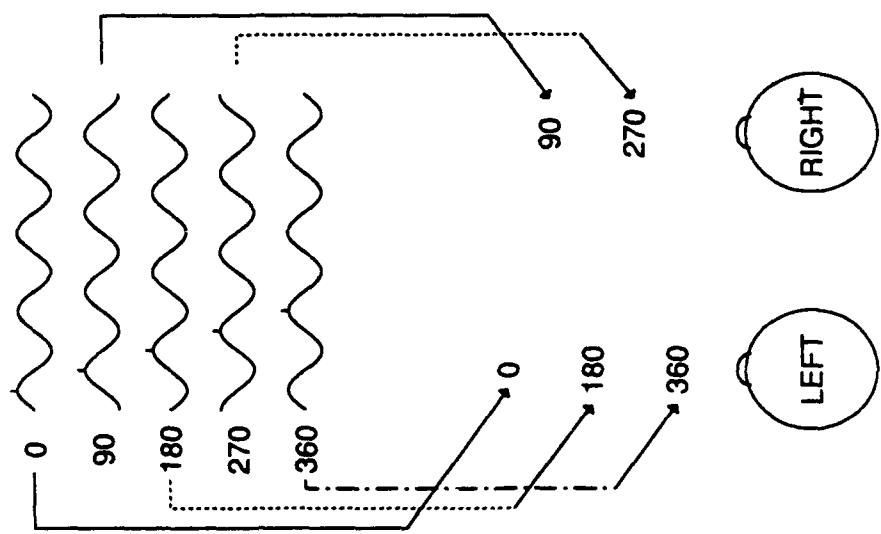


FIG. 2

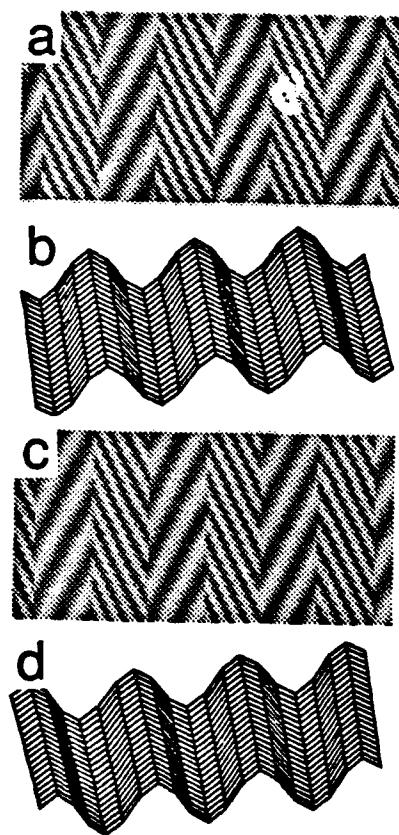


FIG. 3

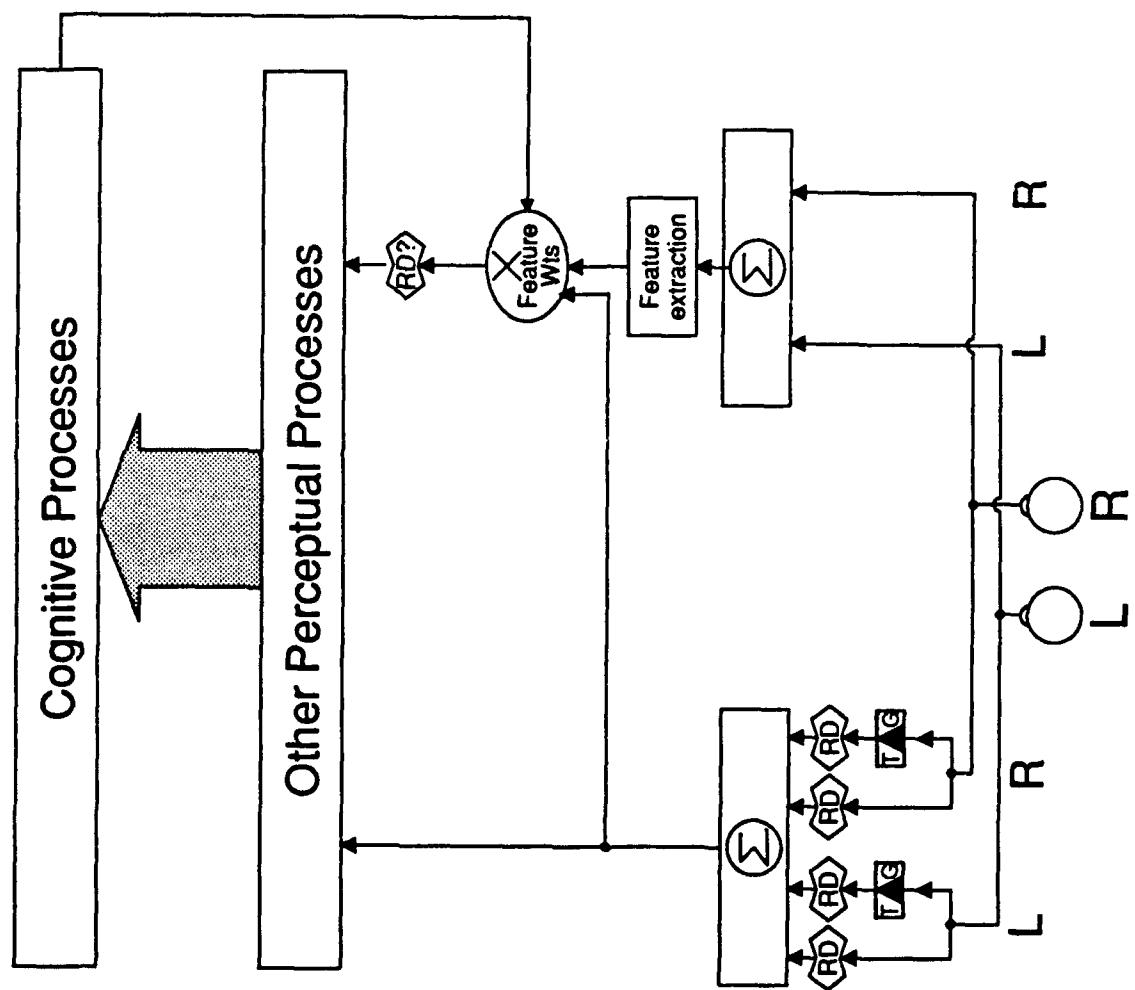


Fig. 4

## George Sperling: HIP Lab Publications, 1993-94

1993 Sperling, G. Wurst, S. A., and Lu, Z-L. (1993). Using repetition detection to define and localize the processes of selective attention. In D. E. Meyer and S. Kornblum (Eds.), *Attention and Performance XIV: Synergies in Experimental Psychology, Artificial Intelligence, and Cognitive Neuroscience - A Silver Jubilee* Cambridge, MA: MIT Press. Pp. 265-298.

1993 Werkhoven, P., Sperling, G., and Chubb, C. (1993). The dimensionality of texture-defined motion: A single channel theory. *Vision Research*, 1993, 33, 463-485.

1993 Solomon, J. A. and Sperling, G. (1993). Fullwave and halfwave rectification in motion perception. *Investigative Ophthalmology and Visual Science*, ARVO Supplement, 1993, 34, No. 4, 976. (Abstract)

1993 Shih, Shui-i and Sperling, G. (1993). Visual search, visual attention, and feature-based stimulus selection. *Investigative Ophthalmology and Visual Science*, ARVO Supplement, 1993, 34, No. 4, 1288. (Abstract)

1993 Lu, Zhong-Lin and Sperling, G. (1993). 2nd-order illusions: Mach bands, Craik—O'Brien—Cornsweet. *Investigative Ophthalmology and Visual Science*, ARVO Supplement, 1993, 34, No. 4, 1289. (Abstract)

1993 Chubb, C., Darcy, J. and Sperling, G. (1993). Metameric matches in the space of textures comprised of small squares with jointly independent intensities. *Investigative Ophthalmology and Visual Science*, ARVO Supplement, 1993, 34, No. 4, 1289. (Abstract)

1993 Sperling, G. (1993). Spatial, Temporal, and Featural Mechanisms of Visual Attention. *Spatial Vision*, 7, 86. (Abstract)

1993 Gegenfurtner, K. and Sperling, G. (1993). Information transfer in iconic memory experiments. *Journal of Experimental Psychology: Human Perception and Performance*, 19, 845-866.

1993 Solomon, J. A., and Sperling, G., and Chubb, C. (1993). The lateral inhibition of perceived contrast is indifferent to on-center/off-center segregation but specific to orientation. *Vision Research*, 33, 2671-2683.

1994 Solomon, Joshua A., and Sperling, George. (1994). Full-wave and half-wave rectification in 2nd-order motion perception. *Vision Research*, 33. (In press.)

1994 Sperling, G., and Dosher, B., A. (1994). Depth from motion. In *Early Vision and Beyond*. Cambridge, MA: MIT Press, 1994. (In press.)

1994 Sperling, G., Chubb, C., Solomon, J. A., and Lu, Z-L. (1994). Fullwave and halfwave processes in second order motion and texture. In *Higher-order processing in the visual system*. Chichester, UK: Wiley (Ciba Foundation Symposium, 184). Pp. (In press.)

1994 Sperling, G., and Lu, Z-L. (1994). Immunity to pedestals distinguishes motion-energy from feature-tracking motion-perception mechanisms. *Investigative Ophthalmology and Visual Science*, ARVO Supplement, 1994, 35, No. 4, 1390. (Abstract)

1994 Sperling, G. (1994). Second-order perception. *Investigative Ophthalmology and Visual Science*, ARVO Supplement, 1994, 35, No. 4, 1477. (Abstract)

1994 Lu, Z-L. and Sperling, G. (1994). Deriving the dimensions of texture perception from metameric texture matches. *Investigative Ophthalmology and Visual Science*, ARVO Supplement, 1994, 35, No. 4, 2161. (Abstract)

1994 Sperling, G., Chubb, C., Solomon, J. A., and Lu, Z-L. (1994). Visual preprocessing: First- and second-order processes in the perception of motion and texture. In *IEEE* (In press.)

1994 Shih, Shui-i and Sperling, G. (1994). Using cluster analysis to discover and characterize covert strategies. *Psychological Science*, 5, 175-178.

1994 Werkhoven, Peter, Sperling, George, and Chubb, Charles. (199x). Perception of apparent motion between dissimilar gratings: Spatiotemporal properties. *Vision Research*, 33. (In press.)

1994 Solomon, J. A., and Sperling, G. (1994). 1st- and 2nd-order motion and texture resolution in central and peripheral vision. *Vision Research*, 33. (6 pp, in press.)

Papers Under Submission for Publication.

199x Lu, Zhong-Lin and Sperling, George. (199x). The functional architecture of human visual motion perception. Submitted for publication.

## **George Sperling: Talks at Symposia and Meetings of Professional Societies: 01Dec93 to 31May94**

† Indicates an invited address.

\* Indicates an abstract of talk was published.

1993 †Sperling, George. International Workshop on Digital Video for Intelligent Systems. Hosted by Department of Electrical and Computer Engineering, University of California, Irvine, California. December 17, 1993. *An Engineering Model of Human Visual Processing / Intelligibility of Extremely Reduced Images*.

1994 †Sperling, George. Western Psychological Association, Kona, Hawaii, Invited Lecture, April, 22, 1994, *Second-Order Perception*.

1994 \*Sperling, G., and Lu, Z-L. Talk presented by G. Sperling. Association for Research in Vision and Ophthalmology, Sarasota, Florida, May 2, 1994. *Immunity to Pedestals Distinguishes Motion-Energy from Feature-Tracking Motion-Perception Mechanisms*.

1994 †\*Sperling, George. Association for Research in Vision and Ophthalmology, Sarasota, Florida, May 2, 1994. *Second-Order Perception*.

1994 \*Lu, Z-L., and Sperling, G. Talk presented by Z-L Lu. Association for Research in Vision and Ophthalmology, Sarasota, Florida, May 6, 1994. *Deriving the Dimensions of Texture Perception from Metameric Texture Matches*.

1994 †\*Sperling, G, Lu, Z-L, Chubb, C, and Solomon, J. A. Talk presented by Z-L Lu. Joint meeting: World Congress of Computational Intelligence/IEEE Neural Networks/Fuzzy Logic/Evolutional Computation. Orlando, Florida. Plenary lecture: May 31, 1994. *Visual Preprocessing: First- and Second-Order Processes in the Perception of Motion and Texture*.

## **George Sperling: Invited Lectures at Universities and Institutes 01Dec93 to 31May94**

1993 University of California, Berkeley. December 3, 1993. *A Theory of Spatial Attention*.

1994 Ohio State University, Columbus, College of Optometry. May 9, 1994:  
Low-Vision Seminar. *Optimal Images and Optimal Perception*  
Distinguished Speakers Series. *Second-Order Perception*